BASETY : Mode d'emploi

Pour ceux qui désirent utiliser BASETY pour construire leur matériel expérimental ou autres merci de citer cet article comme référence :

Léger, L, Boumlak, H & Tijus, C (en révision). BASETY : extension et typicalité pour 21 catégories d'objets. Revue Canadienne de Psychologie Experimentale.

Contact : laure.leger@cognition-usages.org

BASETY disponible en ligne :

http://cognition-usages.org/chart

http://leger.laure.free.fr

Voici un court descriptif de l'article qui permet d'utiliser la BASETY avant la version définitive de l'article.

Résumé

Basety est une base de données qui présente pour 21 catégories d'objets les exemplaires cités par 100 individus natifs français et leur typicalité. Ces 21 catégories sémantiques sont ANIMAUX, ARBRES, ARMES, EDIFICES, FLEURS, FRUITS, INSECTES, INSTRUMENTS DE MUSIQUE, JEUX, JOUETS, LEGUMES, MAMMIFERES, MEUBLES, OISEAUX, OUTILS, POISSONS, PROFESSIONS, RECIPIENTS, SPORTS, VEHICULES, et VETEMENTS. La typicalité de chaque exemplaire a été établie en considérant son nombre d'occurrences dans les cinq premiers rangs de citation. Basety présente différents indices qui permettent de caractériser chaque exemplaire cité : fréquence de citation, degré d'appartenance catégorielle, fréquence lexicale, nombre de

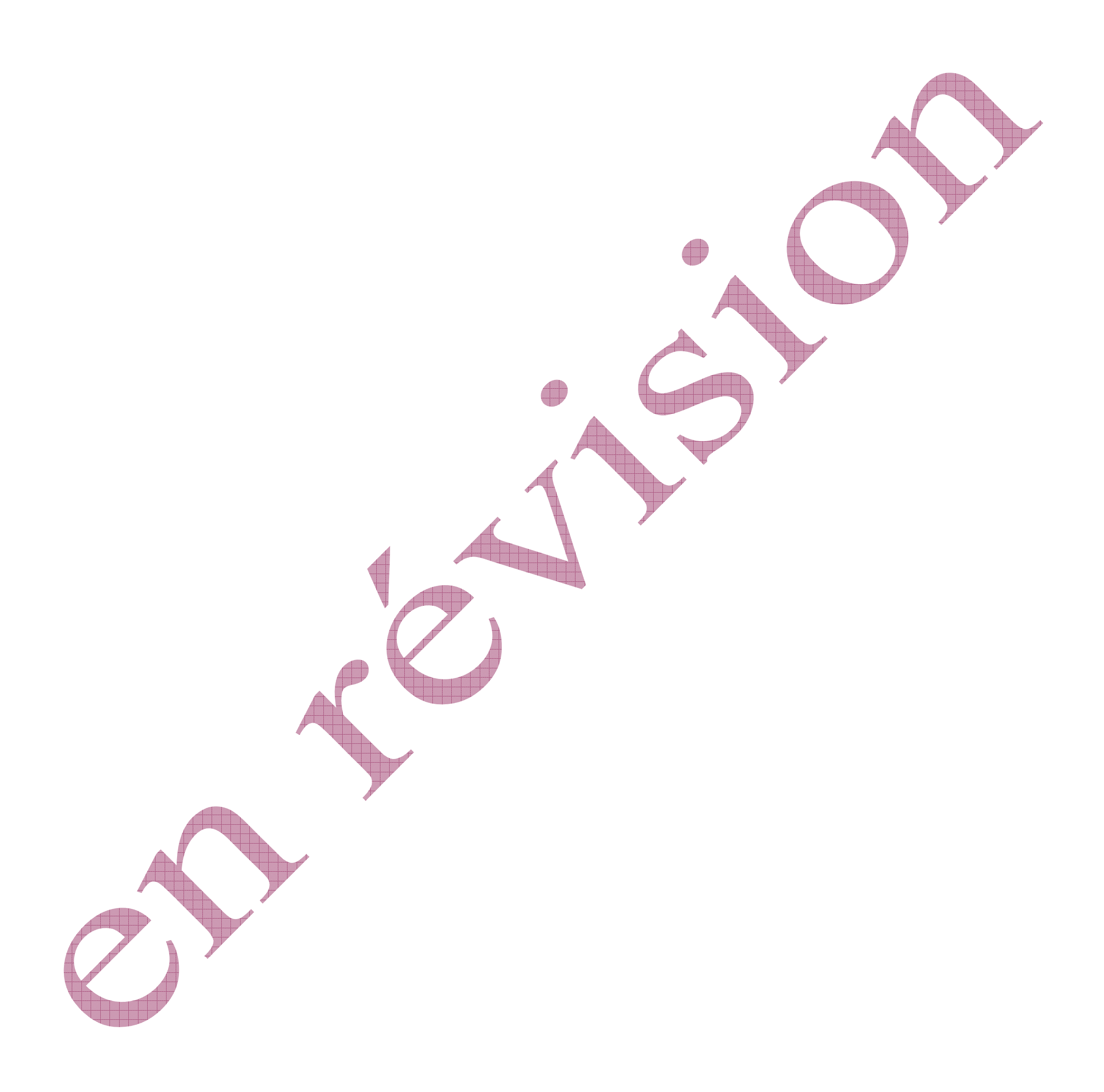
lettres. Pour chaque catégorie nous avons également déterminé leur entropie, leur étendue, leur nombre de prototypes. Basety est consultable à partir de deux sites web : http://www.cognition-usages.org/chart et http://leger.laure.free.fr .

Mots clés : base de données sémantiques, indice de typicalité, catégories sémantiques

Présentation de BASETY

BASETY se présente sous la forme d'une feuille de tableur comprenant plusieurs colonnes. A chaque ligne correspond un exemplaire d'une catégorie. La première colonne désigne la catégorie et la seconde colonne donne les différents exemplaires. Chaque exemplaire est caractérisé par un certain nombre de propriétés, réparties dans les colonnes suivantes. Ces propriétés sont :

- le nombre d'occurrences de l'exemplaire (sur une base de 100 participants) : NOC qui correspond à l'indice de production
- le nombre d'occurrences dans les 5 premiers rang : TYP qui correspond à l'indice de typicalité
- le rang de citation dans les cinq premiers : RANG
- la fréquence lexicale issue de lexique.org : FQLEXlivres et FQLEXfilms
- le fait d'être un mot composé ou pas : C pour mot composé ou S pour mot simple
- le nombre de caractères composant l'item (espace inclus) : NCAR
- le nombre de lettres composant l'item : NLET
- le genre de l'item : GEN ; M pour masculin et F pour féminin
- la fréquence où cet exemplaire a obtenu la note de 5 pour son degré d'appartenance (degré d'appartenance maximale) : AP5
- la fréquence où cet exemplaire a obtenu la note de 4 pour son degré d'appartenance : AP4.


en révision